Accessing online data for youth mental health research

Philos. Technol.
DOI 10.1007/s13347-017-0286-y

RESEARCH ARTICLE

Accessing Online Data for Youth Mental Health Research: Meeting the Ethical Challenges

Elvira Perez Vallejos[1] · Ansgar Koene[2] · Christopher James Carter[3] · Daniel Hunt[4] · Christopher Woodard[5] · Lachlan Urquhart[6] · Aislinn Bergin[7] · Ramona Statache[8]

Received: 12 September 2016 / Accepted: 18 September 2017

Abstract This article addresses the general ethical issues of accessing online personal data for research purposes. The authors discuss the practical aspects of online research with a specific case study that illustrates the ethical challenges encountered when accessing data from Kooth, an online youth web-counselling service. This paper firstly highlights the relevance of a process-based approach to ethics (Markham and Buchanan 2012) when accessing highly sensitive data and then discusses the ethical considerations and potential challenges regarding the accessing of public data from Digital Mental Health (DMH) services. It presents solutions that aim to protect young DMH service users as well as the DMH providers and researchers mining such data. Special consideration is given to service users' expectations of what their data might be used for, as well as their perceptions of whether the data they post is public, private or open.

✉ Elvira Perez Vallejos
elvira.perez@nottingham.ac.uk

Ansgar Koene
ansgar.koene@nottingham.ac.uk

Christopher James Carter
christopher.carter@nottingham.ac.uk

Daniel Hunt
daniel.hunt@nottingham.ac.uk

Christopher Woodard
christopher.woodard@nottingham.ac.uk

Lachlan Urquhart
lachlan.urquhart@nottingham.ac.uk

Aislinn Bergin
a.bergin@chester.ac.uk

Ramona Statache
R.Statache@mmu.ac.uk

Published online: 12 October 2017

We provide recommendations for planning and designing online research that includes vulnerable young people as research participants in an ethical manner. We emphasise the distinction between public, private and open data, which is crucial to comprehend the ethical challenges in accessing DMH data. Among our key recommendations, we foreground the need to consider a collaborative approach with the DMH providers while respecting service users' control over personal data, and we propose the implementation of digital solutions embedded within the platform for explicit opt-out/opt-in recruitment strategies and 'read more' options (Bergin and Harding 2016).

Keywords Ethics · Social media research · Consent · Online data · Data privacy

1 Introduction

Accessing language data from online services can be ethically sensitive, especially when data is derived from Digital Mental Health (DMH) services. This paper presents an overview of ethical issues in online research, with a case study in mental health. This approach provides a focus for discussion on how the research team dealt with the specific ethical issues and controversies encountered.

Young people often experience severe and potentially long-lasting psychological issues, yet many express difficulties in communicating their concerns to professionals. Various studies have indicated that less than 25–35% of those with a diagnosable mental health condition access professional support (Department of Health 2015). Low access to mental healthcare support is becoming a serious problem that is aggravated during the transition from child and adolescent mental health services (CAMHS) to adult mental health services. Consequently, young people are actively seeking alternative and complementary sources for psychological support and online advice (Richwood et al. 2015). From a research perspective, the content posted on Digital Mental Health (DMH) services has the potential to provide a rich, cost-effective, timely and valuable source of data, not least for corpus-based methods, which are well suited to interrogate the volume of text generated online.

[1] Psychiatry and Applied Psychology, NIHR Nottingham Biomedical Research Centre; NIHR MindTech Healthcare Technology Cooperative, Institute of Mental Health, University of Nottingham, Nottingham, England, UK

[2] Horizon Digital Economy Research Institute, University of Nottingham, Nottingham, England, UK

[3] The Haydn Green Institute for Innovation and Entrepreneurship, University of Nottingham, Nottingham, UK

[4] School of English, University of Nottingham, Nottingham, England, UK

[5] Department of Philosophy, University of Nottingham, Nottingham, England, UK

[6] Horizon Centre for Doctoral Training (CDT) and Mixed Reality Lab, University of Nottingham, Nottingham, England, UK

[7] Faculty of Health and Social Care, University of Chester, Chester, England, UK

[8] Manchester Metropolitan University, Manchester, England, UK

Corpus linguistics is the study of linguistic phenomena through large collections of machine-readable text: corpora. The focus of applied linguistic research has always been on real language use; corpus linguistics, in particular, has developed rapidly in the last few years, partly due to the increased possibilities offered by easy access to machine-readable text available in online environments (McEnery and Wilson 2001). As well as allowing researchers to search large amounts of data, corpus programmes have developed new ways of organising and interrogating data. Some of the most popular corpus methods include keyword analysis, concordances and collocates (Adolphs 2006; Baker et al. 2009; Hunt and Harvey 2015). Despite the importance of this form of communication and the potential for its linguistic analysis, collecting and analysing language data from DMH services posted by vulnerable individuals, such as adolescents' expressions and experiences of psychological distress, poses unique ethical challenges that should be carefully considered by researchers.

Only a few peer-reviewed articles have explicitly addressed the ethical dilemmas involved in mental health research when using online data created by vulnerable young adults (e.g., The SharpTalk study on self-harm by Sharkey et al. 2011). The scarcity of ethically sound case studies may be perceived by some researchers as an opportunity to access so-called public data or data within the public domain by registering and becoming part of an online community group designed to support people with mental health problems. Such access to discussion feeds is relatively easy to achieve but carries with it the ethical issues of covert observation, especially when dealing with vulnerable users who post messages for comfort and support and whose potential mental distress at the time of posting might affect their awareness and/or sensibility to being under observation by a 'lurking' researcher. While it could be argued that studies which only passively collect data for linguistic analysis do not require the same ethical considerations as studies that actually involve intervention and manipulation, researchers interested in collecting sensitive data from password-protected DMH services (e.g. depression and anxiety forums) without informed consent should reconsider this poor research practice due to the potential harm to the service provider (e.g. credibility), service user (e.g. violation of trust) and research community in general (e.g. reputation). Equally worrying is the attitude to ethics by editors and reviewers of academic journals who do not perceive the need to develop a culture of respect and trustworthiness and may accept research using online data collection that applies questionable ethical practices. Likewise, DMH services sometimes lack the knowledge required to support decision-making on which research to participate in, and how best to involve their users in this process (Bergin and Harding 2016).

Traditionally, the social sciences have well-established codes of ethics and Institutional Research Boards (IRBs) through which best practice is a negotiated process to ensure research is conducted ethically. However, the involvement of other disciplines in human subject research in recent years has brought with it utilitarian and moral challenges to these established guidelines. These disciplines have not often had to contend with research that involves human subjects and may be inclined to regard it as removed from the individual participant by the calculative nature of the analysis. Consequently, these disciplines may perceive the constraints that seem to be placed on research by ethical reviews as less applicable to their research framework (Metcalf 2016). For example, a researcher may use this type of moral reasoning to justify accessing sensitive data without permission from platform moderators, data controllers

or online users if the potential societal benefits overcome any potential harm caused to others, i.e. the reasoning whereby 'the ends justify the means'. Research driven by data-centric approaches may prioritise securing data access/collection, without major considerations for the context in which the analysis takes place. In other contexts, however, we do not tend to think that the unrestricted pursuit of quick wins or of the greater good is ethical, since this can violate important rights or interests of individuals. Ethical considerations and user-centric approaches should still be among the factors used to target, shape and enhance projects. A user-centric perspective seeks to approach ethical and technological issues from the perspective of users who created the original data and their experiences of the research context. Consequently, we advocate consideration of the ethical issues that arise in this sort of research, adopting a user-centric approach.

From a legal perspective, there is a need to balance the end goals of research projects that aim at the greater good against the regulatory context to ensure that virtuous research goals do not override legal safeguards. Therefore, the utilitarian goals of research do not absolve studies from legal compliance requirements. This is especially so for privacy and data protection considerations, as these are fundamental, inalienable human rights (Article 8 of the European Convention on Human Rights and Articles 7 and 8 of the EU Charter of Fundamental Rights) which often enable other human values like dignity, autonomy or identity formation. When dealing with vulnerable populations including children and young people, the need to respect these elements increases. This becomes problematic, however, when handling big social media data for research because of significant technical and regulatory issues. The 'velocity, variety and volume' of big data is also a challenge computationally (ICO 2014a, p. 6). These large datasets often involve personal data, and accordingly this brings the full force of EU data protection regulations to the fore. The UK Information Commissioner's Office (ICO 2014b) has many concerns around use of big data analytics with personal data, yet they state 'big data is not a game that is played by different rules' and the existing data privacy rules are fit for regulation (ICO 2014a, p. 4). They are particularly concerned about a number of issues, such as the following: ensuring sufficient transparency and openness with data subjects about how their data is used; reflecting on when data repurposing is or is not compatible with the original purposes of collection (e.g. data collected for one purpose is reused for another); the importance of privacy impact assessments; and how big data challenges the principle of data minimisation and preservation of data subjects' access rights (ICO 2014a, p. 5–6). The legal aspects of privacy online are complex, especially when considering the global nature of the internet and the storage of data in servers located in different countries with different data protection laws. The EU has passed the General Data Protection Regulation (GDPR) 2016, to be enforced across all EU member states from May 2018. Drafted with the digital age in mind, unlike the former Data Protection Directive 1995 (DPD 1995), it expands on the old law by providing a wide range of data subject rights. These will impact on abilities to do big social media research as users have a right to data portability (to receive their data from a service, and to transmit to another) to the right to object to processing too (http://ec.europa.eu/justice/data-protection/). Furthermore, the landscape around international data flows outside of the EU is shifting. The global scope of social media services coupled with regional jurisdictional rules means the landscape for researchers outside the EU legally accessing data involving EU citizens could become more complex. To explain,

Article 3(2) GDPR expands the scope of EU DP law to data controllers targeting goods or services towards EU citizens, who need to comply to gain market access. With use of services hosted in the cloud, establishing which jurisdiction they are based in is increasingly important if they are geared towards EU citizens. Furthermore, with data flowing outside of the EU, the destination country needs to be deemed to provide adequate protections. With the US case, the Safe-Harbour Agreement (which collapsed at the end of 2015) existed to provide this assurance, and has since been replaced by another bilateral EU-US agreement, the so-called Privacy Shield, to establish rules on US-EU data transfer, including safeguards. Important for UK researchers, with the changing UK-EU relationship due to Brexit, the adequacy of the UK DP framework post Brexit, even if based on GDPR (Denham 2016), remains to be seen (partly due to recently passed UK surveillance laws) (Edwards 2016). This means a similar agreement for UK-EU data transfers may be necessary in the future. This all adds to complexities of doing research in compliance with law on social media data from users around the world, hosted in different locations globally.

Moreover, the infancy of internet research is aggravated by a lack of consensual institutional guidance (e.g. differences between schools from the same university) and formal training for those academics involved in social media research (Carter et al. 2015). Therefore, the wide variety of academic attitudes and lack of consensus towards the ethical challenges of online research is not surprising. Several sets of ethical guidelines and recommendations for internet research have recently emerged (e.g. Ess 2002; BPS 2013), identifying some of the key ethical issues relevant for social media researchers. Access to these documents, however, does not necessarily imply compliance and consistent interpretation, especially when considering the myriad social media contexts in which these guidelines would need to be applied. The ethical principles discussed in this paper are mainly derived from the Association of Internet Researchers (AoIR) Ethics Guide (Markham and Buchanan 2012), Recommendations on Good Practice in Applied Linguistics (British Association for Applied Linguistics 2006) and the 'Online survey tools: ethical and methodological concerns of human research ethics committees' (Buchanan and Hvizdak 2009). Markham and Buchanan (2012) recognise the need for further comment and negotiation of online guides to ethical research involving human subjects as evolving technologies and software present new ethical challenges. This paper hopes to outline the challenges, and suggest solutions, to issues related to linguistic analysis of online mental health communities, recognising that young people often utilise these forums (Burns et al. 2010), contributing to the debate regarding online sensitive research.

As outlined by Markham and Buchanan (2012), 'the basic tenets shared by these policies include the fundamental rights of human dignity, autonomy, protection, safety, maximisation of benefits and minimisation of harms, or, in the most recent accepted phrasing, respect for persons, justice and beneficence'. (p. 4). These principles are further instantiated through discipline-based guidelines including the Association for Computing Machinery's (ACM) 'Code of Ethics and Professional Conduct' (Anderson 1992), Social Media & Social Science Research Ethics (Townsend and Wallace 2016) and the British Psychological Society's (BPS 2014) 'Code of Human Research Ethics', which particularly emphasises the personal and professional responsibilities of researchers.

The Code of Human Research Ethics (BPS 2014) outlines four main principles underpinning the ethical conduct of research: (1) Respect for the autonomy and dignity of persons, (2) scientific value, (3) social responsibility and (4) maximising benefits and minimising harm. As outlined in these various guidelines, the following issues often pose particular challenges when evaluating the ethics of (proposed) internet-based study:

1. Public-private domain distinction online
2. Confidentiality and security of online data
3. Procedures for obtaining valid consent
4. Procedures for ensuring withdrawal rights and debriefing
5. Implications for scientific value and potential harm

Using these issues as a foundation, this paper explores the practical application of key ethical principles for online data access from a user-centric perspective and identifies the ethical issues that have to be surmounted during the research planning stages. To contextualise key issues, we will present a specific case study and consider the ethical challenges encountered when accessing data from Kooth, a UK-based online counselling service designed to provide community as well as professional support to children and young people.

1.1 Our Case Study

Kooth is an online counselling and emotional well-being platform for children and young people, accessible through mobile, tablet and desktop and free at the point of use. Kooth was established in 2004 and it is part of the XenZone family, a provider of online mental health services for children, young people and adults established in 2001. XenZone practitioners are Organisational Members of the BACP (British Association of Counsellors and Psychotherapists) and all clinical staff hold memberships with the various bodies that monitor the counselling and psychotherapy professions, such as the United Kingdom Council for Psychotherapy, the Health Professions Council and the BACP. Their vision is to lead the way in using digital technology to remove barriers to achieving emotional and mental health. XenZone works as part of comprehensive CAMH (Children and Adolescents Mental Health) services by providing an early intervention service through their online counselling and emotional well-being support platform Kooth. When commissioned by local authorities, any child or young person can access Kooth for free. Where a young person meets the criteria for specialist CAMH services, Kooth works with the local NHS partners to ensure an integrated model of support. XenZone also develops online resources (e.g. Online Risk and Resilience resource) in partnership with other institutions including the Department of Health.

Kooth provides different modalities for communicating with their service users. It includes private one-to-one communication with a qualified counsellor (higher expectation of online privacy), as well as support from the Kooth community (peer support open forum, lower expectation of online privacy). From a user-driven perspective, we are interested in accessing censored posts only identified within the communications taking place at the peer support open forum. We understand censored posts to mean any

communications that are intercepted automatically or manually by the forum moderator because of inappropriate content. All these data will be analysed applying content analysis and corpus linguistic approaches. A close linguistic analysis will characterise the content of the posts through keyword analysis to understand if the posts contain sensitive/inappropriate information that would justify their removal and how they might have been identified. This analysis will provide both a content analysis and a quantitative analysis (keyword analysis) of the frequency and nature of censored texts. This information will bring insights into users' online behaviour and the function of the DMH platform, for example, how the existing moderation protocol reflects the codes of practice set out by the platform providers. The analysis will also consider whether the current moderation process captures all of those potentially sensitive posts (i.e. sensitivity and specificity). The analysis will also consider the linguistic context of the censored text (i.e. conversation analysis within models of computer-mediated discourse analysis of the communications that preceded and followed it as well as interaction patterns). Censored posts can include inappropriate language and distressing information but also requests for disclosure of personal information from other forum users. They can also contain information that is not suitable for the forum, such as lengthy personal statements which may be more suitable for the in-house blog section, rather than the live peer-to-peer forum, or the unintentional disclosure of users' personal information (e.g. date of birth) that could jeopardise service users' confidentiality agreements. While representing a substantial dataset, censored posts from Kooth have never been systematically analysed. Results could bring insights to support the improvement of current automatic recognition and identification systems of inappropriate content, as well as guidance on how moderators should deal with such content (e.g. legal and ethical responsibilities). For example, posts that contain swear words are often automatically deleted; however, these posts could also contain distressing information that could flag risks to moderators (Stern 2003). The following sections of this paper will focus on the multiple ethical issues presented by accessing and collecting data in this context; a subsequent study will focus on the linguistic analysis to interpret and contextualise the findings.

2 User-Centric Approaches to Ethics: Planning and Designing Social Media Research

In this section, we discuss the potential benefits that user-centric approaches could have on the planning and design of social media research. Some of the complicating factors when considering the ethics of using social media data reside in understanding the users' perspective, their interpretation of online privacy and sense of personal ownership of data. The online public-private domain distinction and the related differences between open data vs. public data are relevant for managing DMH users' expectations of privacy as opposed to the expectations of privacy held by the platform provider or the researcher. Users' perceptions of privacy can also influence perceptions of confidentiality and security and these would vary depending on the medium used for communication, with one-to-one communication perceived as more private, confidential and secure than a peer-to-peer open forum. We argue that user-centric

approaches to obtain valid consent and inform users about their right to withdrawal and debriefing have the potential to minimise harm and add scientific rigour to online research involving mental health data.

2.1 Online Privacy and Data Ownership

On the topic of online privacy, the AoIR documents refer to the definition provided by the American Civil Liberties Union which puts it as, '[the ability] to ensure that individuals have control over their data' (American Civil Liberties Union 2015). Thus privacy is not necessarily about making sure that no one has access to the data, but rather about having the ability to control who does and does not have access. The way in which privacy is surrendered when an individual makes an informed decision to consent to the publishing of data on potentially open forums such as social media sites thus stems from their limited ability to control access to the data by other site users and organisations whenever they desire to do so. Privacy concerns, however, are not limited to control over personal information. Helen Nissenbaum (2015) proposes a more user-centred framework to understand privacy issues by pointing to context, social norms, expectations and values as factors influencing the distinction between public and private. For example, a factor determining whether a privacy violation has occurred or not relates to societal standards of intimacy and the degree of sensitive or confidential information disclosed. Accordingly, members of online forums disclosing intimate information (e.g. www.silentsecret.com) will expect higher degrees of privacy and will be more sensitive to intrusion than others on forums where non-intimate data is discussed. Similarly, forums that are password-protected can be perceived as more private than those that are password-free. For example, Robinson (2001) identifies passwords as 'gatekeepers' to information, requiring explicit consent to be obtained for access to data.

Our privacy, however, is increasingly networked and control over individual personal data can be very difficult as data in networked contexts is circulated, tagged, shared and aggregated in unpredictable ways (Boyd 2012; Boyd and Crawford 2012). In the case of Kooth information posted by teens is often intended only for a specific networked public made up of peers, a support network or specific community, not necessarily the public at large (Boyd 2014). The lack of perceived control over privacy and information flow within networked contexts is inevitably influencing users' expectations. According to the theory of contextual integrity (Nissenbaum 2004), the online user has the right to know who is gathering their personal information, who is analysing it and for what purpose, who is disseminating it and to whom. From a user-centric perspective, ordinary site users may not even know their data is being used, making it very hard for them to protect their own rights. Accordingly, it is important to ask whether the information can harm users; interfere with their self-determination; or amplify inequalities, stigma, status or lack of credibility.

Personal ownership of data in this context relates to control over the act of passing the data on to other people/organisations and the use of the data for further purposes (e.g. research). This may include concerns where DMH site users' contributions are shared between researchers and also where data is uploaded for analysis on third party servers (such as WMatrix (http://ucrel.lancs.ac.uk/wmatrix/)). As with online privacy, users may harbour strong

feelings of ownership of the information posted on social media, particularly when the information is descriptive of users, and it is personal and intimate. These subjective feelings about data ownership may hold regardless of legal issues of intellectual property and data ownership determined by the platform. It is important to clarify that while data can be owned in the USA, and therefore be traded, this goes against the European approach of privacy as a human right.

2.2 Open Data vs. Public Data

The ill-fated Samaritans Radar app provides a striking example of the difficulty in distinguishing between private and public communication, as well as the legal and ethical consequences associated with this (Perez Vallejos 2015). The Radar app used linguistic analysis and the Twitter API in an attempt to identify tweets that could be associated with suicidal ideation. When certain trigger words or phrases were tweeted (i.e. keyword identification associated with suicidal ideation), the app would alert the friends of the person tweeting. Following public outcry against perceived privacy invasion by the app, the ICO launched an inquiry which ultimately found the Radar app to be not only ethically unacceptable but also legally in violation of data protection laws. Two key elements were identified by the ICO; first, that data protection principles could be breached even when the information was obtained from a publicly available source and, second, that data protection implications must be considered if analytics are used to make automated decisions that could have a direct effect on individuals. Moreover, the app was criticised for generating false alarms when keywords were taken out of context and the general issue of identifying authorship which runs counter to anonymisation arguments.

Anonymisation is one of the most basic steps for maintaining confidentiality and is also recommended, where possible, by the Data Protection Act 1998 when dealing with personal data. That being said, the ICO argues that, 'anonymisation should not be seen merely as a means of reducing a regulatory burden by taking the processing outside the Data Protection Act 1998. It is a means of mitigating the risk of inadvertent disclosure or loss of personal data, and so is a tool that assists big data analytics and helps the organisation to carry on its research or develop its products and services.' (ICO 2014a para. 46). The need to protect the anonymity of participants is even more relevant when research uses data from online sources where access to the raw data cannot be controlled by the researcher. The wealth of secondary information sources linked to a particular individual is making it increasingly easy to de-anonymise data by combining and aggregating publicly available personal data. For example, when reporting research findings direct quotes from posts can easily be traced back to identifiable individuals, and even when personal information is omitted or altered, researchers should always consider the likelihood of re-identification.

The distinction between open data and public data is crucial to understanding the inappropriateness of accessing and collecting online data from public forums without explicit consent from users or data controllers (e.g. platform moderator). Open data is explicitly intended for a public audience without expectations of privacy. However, public data do not necessarily fall within that category. To quote the Open Data Institute (ODI), 'Not all data is open data […] Open data is data that anyone can access, use and share […] Open data is data that is published under a licence with express permission to

reuse, share and modify'. In contrast, from a strictly legal perspective, only documents that are not protected by copyright law should be classed as being 'in the public domain'. This may be because the creator has given unrestricted access via, for example, a Creative Commons licence (Erickson et al. 2015) or the documents do not qualify for copyright. In the UK, there are certain exceptions, including the use of material for 'data mining'. More information about exceptions can be found here: https://www.gov.uk/guidance/exceptions-to-copyright. Based on these definitions, we can deduce that while all open data is public data, not all public data can be considered as open data and therefore obtaining consent is recommended when accessing public online data. More broadly, human rights law states that individuals have a reasonable expectation to privacy, even in public spaces (Von Hannover v Germany 2005).

Clearly, legally and ethically, determining whether a data set is of an open or private nature has far reaching consequences for the types of restrictions that apply to the way in which data is managed, analysed and reported. Unfortunately, when dealing with data from online sources, e.g. discussions on user-groups or social networks, the distinction between private and public can frequently be a challenging process. When deciding how to deal with data from the internet, the terms and conditions of the online platform can only ever serve as an initial starting point in determining the publicness of the data. As stated by the AoIR Ethics working committee, 'People may operate in public spaces but maintain strong perceptions or expectations of privacy. They may acknowledge that the substance of their communication is public, but demand that the specific context in which it appears implies restrictions on how that information is – or ought to be – used by other parties. Social, academic, or regulatory delineations of public and private as a clearly recognisable binary no longer holds in everyday practice.' Some of the reasons why people may consider their publicly accessible internet activity to be private, despite agreeing to the site User Licence Agreement, are related to the fact that communication on the Internet has important characteristics of persistence (i.e. online expressions are automatically recorded and archived), replicability (i.e. content made out of bits can be duplicated), scalability (i.e. the potential visibility of content in networked publics is great) and searchability (i.e. content in networked publics can be accessed through search engines) that do not apply in face-to-face or phone communication (Boyd 2014). People therefore do not have an intuitive sense about the level of privacy they should expect from internet communication.

2.3 Confidentiality and Security of Online Data

Anonymisation is one of the most basic steps for maintaining confidentiality, showing respect and thus gaining the trust of research participants. The UK ICO define anonymization as 'the process of rendering data into a form which does not identify individuals and where identification is not likely to take place' (ICO 2012, Appendix 2). The need to protect the anonymity of participants is even more pressing in research on social media posts where access to the raw data, i.e. the online posts, cannot be controlled by the researcher. At the same time, the wealth of secondary information sources that can be mined in connection to any hint at the identity of a participant is making it increasingly easy to identify data. The classic example of this is the de-anonymisation of users in the AOL Search Log by journalists of the *New York Times* in 2006 (Barbaro and Zeller 2006). In this case, AOL had released a list of 20 million Web

search queries, where the identity of the AOL users was anonymised by replacing their names with 'user numbers'. However, the *New York Times* reporters showed that they were able to relatively easily re-identify individual users by correlating the content of the searches against publically available data sources.

The researcher is responsible for protecting the identity of participants throughout the lifecycle of the research project, even if the research participants are not concerned about data disclosure. To achieve this, we recommend not collecting more data than needed (i.e. data minimisation, one of the principles of Data Protection law), and taking all necessary steps to ensure any personal information is safely secured and managed in a way that ensures anonymity—for example, by ensuring that participants are not identifiable in the outputs of research. Clear and transparent procedures should be developed to protect identities of those who could be identified through third parties. Personal data/identifiers should be kept secure and separate from the individual responses. When using anonymous data for secondary analysis, particular care must be taken to ensure that the subsequent analysis is legitimate (i.e. it is declared on the research protocol) and it retains the anonymity of respondents. A clear timeline for destroying the data in an appropriate timeframe after publication should be ensured. In dealing with personal data, it is important to remember that there can be legal, as well as ethical, requirements for anonymisation. In the case of the UK, the legal requirements are summarised in the UK Data Protection Act 1998, and monitored by the Information Commissioner's Office (ICO). In recognition of the complexity of data anonymisation, the ICO provides a useful code of practice on anonymisation (ICO 2014b). As part of the drive towards 'open data', further advice on anonymisation is increasingly being made available by national data services such as the UK Data Service & UK Data Archive (Corti et al. 2014).

It is important to remember limitations of anonymisation too. The robustness of anonymisation techniques, given their prominence as a policy tool, have been questioned because the dominant 'release and forget model' (i.e. suppression or aggregation of identifiers before public release) can be relatively simply de-anonymised (Ohm 2010). Similarly, the Article 29 Working Party found there are many ways to de-anonymise including 'singling out' (where someone can be picked out of records), 'linkability' (where two records can be put together to re-identify) and 'inference' (where deduction and probabilities are used to on various attributes) (A29 WP 2014, p. 11–12). In seeking to overcome some of these challenges, the UK ICO advise only publicly releasing *necessary* anonymised data because it is hard to know what data is already in the public domain and how it may correspond to new anonymous datasets (ICO 2012, p. 19).

Geolocation data could implicitly build up a very detailed picture of some user's life, even if the data is anonymised in terms of their name. It is important to consider whether inclusion of user location data combined with specific post text is enough to theoretically enable user identification through searches on the platform or other publically available data sources. Potential data linkage should be considered in this context. Location data could be aggregated to a less identifiable level such as region. Presenting such information alongside sentiment scores (i.e. positive or negative tone) for collated data should present lower risk. IP addresses are more widely available and should be separated from user data to ensure anonymisation.

The researcher should also consider how cross-border data will be handled if IP addresses are considered by one country to fall under privacy regulations, or where the servers are housed (e.g. under EU law and regulations). In our case study, service users access Kooth anonymously and therefore there is no need to remove names or email address from a users' posts. If censored posts contain personal information, this would have had to be removed—ideally by the platform—to avoid users being identified. We agreed with the data controllers that only data from UK-based IP addresses would be released to avoid issues related to different data protection regulations between the EU, UK and USA. It is important to also request IP addresses to be removed from the raw data once UK-based IP addresses have been identify to clear any personally identifiable data. Moreover, the corpus linguistic analysis planned to provide quantitative results (e.g. word frequency) will be reassessed to prevent user identification. Due to the sensitivity of the content, direct quotes extracted from censored posts that have never reached the forum (unlikely to be indexed and therefore unlikely to be searchable) will not be published or disseminated to ensure users are protected from being identified by each other. Moreover, results from the content analysis will be presented in a way that preserves user anonymity, for example, by paraphrasing or changing locations or identifiable data.

2.4 Procedures for Obtaining Valid Consent

Valid consent fundamentally deals with respect for the autonomy and dignity of persons. In order for valid consent to take place, it is necessary that the participant is fully aware and has a true understanding of that which is being consented to. This is why, for instance, research involving children requires consent from their legal guardian. Social media research should ensure that any participant involved in primary data collection must freely, and with the appropriate knowledge of the research, give consent to take part.

Legally speaking, using social media data about health requires explicit consent because it is sensitive personal data. Such consent must be an unambiguous 'freely given specific and informed indication of his wishes by which the data subject signifies his agreement to personal data relating to him being processed' (EU Data Protection Directive 1995). The nature of explicitness is not defined in law, but 'it could include a handwritten signature affixed at the bottom of a paper form, but also oral statements to signify agreement, or a behaviour from which consent can be reasonably concluded' (Article 29 Working Party 2011 p. 11). Relying on contract terms and conditions is not likely to meet these thresholds.

A key element in this consideration is the aspect of appropriate knowledge, or 'informedness', in the consent for participation in a study, especially for research that uses accessible social network posts. While this may present a technical and potentially labour-intensive challenge, the advantage of consistently making the extra effort of obtaining valid consent will go a long way towards establishing a conscientious and trustworthy reputation. An illustrative example of potential problems that can arise when proper informed consent is not obtained was provided by the controversy around the Kramer et al. (2014) 'Emotional Contagion through Social Networks' publication. Kramer et al. (2014) asserted that participants had provided consent for the study since 'it was consistent with Facebook's Data Use Policy, to which all users agree prior to

creating an account on Facebook, constituting informed consent for this research'. Specifically, the research—a joint collaboration between researchers from Facebook, Cornell University and the University of California-San Francisco—used an experimental design to investigate the transference of emotional states on Facebook, covertly manipulating the presentation of status updates conveying positive and negative affect that almost 690,000 users would receive within their profile's Newsfeed over the period of one week. This piece of research did not consider that users with mental health issues and low mood could be present in their user sample, causing anxiety and anger among Facebook users. Consequently, the strong reactions among users have set a precedent for influencing and hopefully improving the internal review practices of research groups and online platforms (Kramer 2014). The Data Use Policy however, even if it was actually read by a Facebook user, does not provide any information about the nature of the specific Kramer et al. study. The requirements for full awareness and true understanding of the matter that was being consented to were therefore clearly violated. The statement regarding the implied consent to taking part in the study was especially highlighted in the subsequent 'Editorial Expression of Concern' that the Editor-in-Chief of PNAS published following the negative public response to the paper (Verma 2014).

An example of good practice, however, is the SharpTalk study by Sharkey et al. (2011), in which a negotiated approach to consent (Berry 2004) was considered. SharpTalk was an experimental online self-harm support group designed only for research purposes. The objective of the online group was 'to observe and assess engagement and shared learning between young people that self-harm, health professionals and health care students' (p. 754). Research data included all posts and discussions during the 14 weeks that the forum was open. SharpTalk recruited 77 young people (16–25 years old) who self-harm through existing online self-harm websites. The forum was moderated by 18 NHS professionals and healthcare students that provided support in time of crisis to the participants. Consent as a process provided opportunities for participants to give consent at different times and for different levels of participation (e.g. study registration, participation and use of quotes). Participants chose a unique email address and username to register and participate in the forum.

Consent should be given for what will be undertaken and if, for example, the nature of a participant's involvement changes, they must be informed and, if necessary, consent must be sought again. There should be evidence that consent has been given, at least including the information that research participants were given about the nature and purpose of the study, that consent is voluntary and that they can withdraw at any time. While participants should not be overwhelmed with information, it should be comprehensive enough to allow informed consent and any suitable adjustments to the information should be made available when appropriate (e.g. material produced in minority ethnic languages or large print). As much as possible, participant documentation (e.g. participant information sheet) should be co-produced with a representative sample of the targeted population to ensure documents are clear and accessible. Parsons (2015) pointed out that digital technologies offer an unprecedented opportunity for transforming and supporting informed consent practices for/with children and young people in social research. Taking into account that 14–15 year olds have the highest level of technological knowledge and confidence in digital technologies (Ofcom 2014), it makes sense to support children's understanding of the research process by tailoring methods and information appropriately (ESRC 2015) and consider the importance of

consent as a process rather than a one-off 'tick box' exercise (Dockett and Perry 2011; Clarke and Abbott 2015) by providing child-friendly features and accessible formats such as appropriate and clear language, audio options, interactive interfaces, incorporating images and animations (e.g. comic/cartoon style) (Alderson and Morrow 2004; Tait et al. 2007).

Exceptions for gaining informed consent include covert research, which must be subject to an independent ethical review and legal advice prior to the start of the research project. In certain circumstances, consent can also be gained from a gatekeeper or proxy (e.g. parents). When minors[1] (i.e. under 18s) are involved in research, parental consent from those with a duty of care (e.g. teachers, parents) is mandatory, along with active consent from the minor. This dual consent can often be complex and time consuming but is a necessary response to legislation constraints and the need to ensure that participants are fully informed and understand the nature, purpose and outcomes of the research (Tisdall et al. 2009).

In the case of smaller DMH sites, explicit consent can be sought from both ordinary users and from the moderator or data controller who generally owns the data generated within their platform. In our case, attendance at an academic and professional workshop provided initial access to the Service Development Manager of Kooth. Subsequent email and phone contact was crucial for building trust and the opportunity to negotiate the conditions of this collaborative work. Data controllers requested more information including the purpose of the study, suggestions for data collection, data analysis and dissemination plans. Data controllers were concerned about jeopardising the intimacy and privacy of the peer-to-professional forum. Data controllers opposed including users under the age of 16 due to mandatory parental consent that could breach the confidentiality principles stated in their 'House Rules'. While building a positive relationship with the data controllers was a prerequisite for data access, this was also time consuming and required resources to be allocated and planned in advance to ensure this relationship was not tokenistic. According to our experience, correspondence with larger organisations by researchers is likely to be ignored if an initial approach is not made face to face or over the phone via existing contacts. Therefore, while this method was successful in our own case, we acknowledge that it may be more challenging for researchers studying interactions on large commercial platforms (e.g. Instagram) where efforts to contact data controllers may well be ignored.

If the data is large scale, researchers may consider it impractical to gain consent from each individual and should take into account other ways of informing participants and how that would impact on data use and data representativeness. According to the terms and conditions of Kooth, explicit opt-in consent was not required as it was stated that personal data could be used to improve the quality of the service. Our research team reasoned with the data controllers about the ethical approach we intended to take regarding data collection and informed the user about the use of their personal data for research purposes. It was agreed that a pop-up window would appear on one occasion only for users over 16 to decide to opt-in to the study; a downloadable 'Read

[1] Under the Clinical Trials Regulations (Medicines and Health roducts Regulatory Agency (MHRA, 2004) a minor is a person under 16 years, if these regulations do not apply, the age of majority is 18. For children under 16, who are considered to have the capacity to consent, their consent is valid and the consent of someone with parental responsibility is not required, although it is good practice to gain their agreement.

more' option, previously co-produced with young people to ensure accessible language, was included on this window to inform users about the nature of the study, objectives, dissemination plans, anonymisation, right to withdraw, research team, etc. It was agreed that the research team would cover the software development costs involved in designing and testing such an option. We are aware that this solution may reduce recruitment figures (Hudson and Bruckman 2004) and therefore increase the recruitment period. It was also agreed that site users' feedback would be closely monitored to detect any complaints made to the platform in relation to the pop-up window. Even though Kooth has gathered data from over 12 years of online counselling services, we agreed that only data from users on the peer-support open forum that provided consent via the pop-up window would be collected for a period of 12 months and transferred to the research team via an encrypted memory device.

2.5 Procedures for Ensuring Withdrawal Rights and Debriefing

The right of the individual to withdraw from the study should hold just as much online as it does in offline research. The right to withdraw and the provision of adequate debriefing are both closely linked to valid consent. Since the act of participation often provides a deeper understanding of the true nature of a study, the right to withdraw supports the validity of the consent provided by the participants who remain in the study. The same is true of the debriefing, especially for research where the nature of the study requires that the participants must be naïve to the true purpose of the task/manipulations. One of the challenges for internet-mediated research is the indirect or remote interaction with participants which includes the possibility of participants disappearing from the study (e.g. closing a web browser page of an online questionnaire) without communicating whether they wish their data to be removed and without paying attention to debriefing information. In our case study, users were given the option to withdraw from the study at any time by contacting the platform team via the 'get in touch' email provided within the platform.

Many DMH platforms have their own privacy settings or terms and conditions (T&Cs) for how users' data can be accessed/used (e.g. if they can opt-out of third-party information sharing), which may be relevant for researchers. Most users, however, admit to not reading all the fine print included in terms and conditions and of those who do, only a small minority say they understand it (see Steinfeld 2016 for a review). T&Cs should be informative, clear and easy to read. Several platforms such as Elefriends (managed by Mind, the mental health charity) have included recommendations for clarifying privacy policies and encouraging users to read them. These include presenting the policy in a multi-layer format or adding 'House Rules', a more accessible document to inform users about what to expect from the site, the 'handlers' (i.e. site moderators) and fellow Elefriends. Their T&Cs aim for increased clarity and brevity, using bullet points and summaries of the issues they perceive as most deserving of full attention.

Unfortunately, T&Cs differ for each platform and are subject to regular changes. If T&Cs change during the data collection period of a research project, actions will depend on how data will be used (i.e. aggregated vs. non-aggregated data) and how significant the changes are. Again, a reference group could be useful to consult with participants if they feel their data is being used in a way which they did not agree to.

Informed consent could therefore be sought again. Moreover, when platforms are monitored for prolonged periods of time (e.g. > 24 months), participants may forget that their data is being analysed and a reminder could be recommended. In our case study, only personal data collected and stored retrospectively until the point of consent was collected, therefore there was no need to re-contact users.

It is good practice to provide a forum for participants to ask questions online before consenting to participate in a research project and provide an online contact point for any enquires about the research and a means of raising concerns they may have about the research process. For example, the platform Elefriends provides a feedback page and an email address for users to get in touch. In our case study, we provided a contact email address for users to contact the research team directly. This information was contained within the 'Read more' option presented on the pop up window or by request by contacting the platform team.

2.6 Implications for Scientific Value and Potential Harm

Use of DMH services requires internet access which presents potential barriers in terms of costs, access and usability. This may cause some groups of people to be excluded by virtue of not being able to afford internet access. Similar exclusion risks may apply to people living in parts of the country that offer limited or no internet access. In addition, it is known that internet use is currently more common among younger age cohorts (Office for National Statistics 2014). All these factors and others such as health and disability challenges can create barriers to participation either in themselves or in combination.

It is important to keep in mind that some users may post more than others, therefore frequency and intervals should be taken into account in the methodology. Moreover, when accessing platforms dedicated to vulnerable groups, such as mental health forums, researchers may consider being clear to all platform users from the outset of the engagement that they are being asked to participate in a study, that the researchers are not part of the actual platform user population and that the research is independent of the platform administration/moderation. Seeking advice from a reference group representing that community on any considerations relevant to the research conducted could minimise any potential disruption.

Researchers should be aware that any interaction with active as well as potential participants within a DMH research project could also affect other individuals whom they know. Emotional harm to users and their relatives may be caused through disclosure of identity or misuse of information; data management may be particularly pertinent when informed consent is pending or inferred from the T&C agreed. Individuals may feel harmed if data they had intended to be private is used for research purposes without their explicit knowledge. Social harm can occur when the ramifications of a research project extend beyond the participant themselves and others feel as if their data has been wrongly used. Equally important is the need to protect researchers and the potential harm that the research can have on them, for example, emotional intensity in the early stages of research or being a researcher practitioner (Kidd and Finlayson 2006).

The nature of DMH services means that the data is often already accessible, so consent is sought after the posts are written. Unlike other forms of social research such as interviews and surveys, users may not be aware of how their data might be used until after it was written and posted. Not requesting explicit consent has the potential to feel

like a harmful intrusion to some users whose highly sensitive interactions are being considered for use in a research project. Part of the harm-minimisation process should include communicating to concerned users exactly how their data has or has not been used and for what purpose. Communicating with potential participants via social media also risks third parties viewing the communication. For example, abusive parents may intercept publically available communications between their child and a researcher interested in conducting research on child neglect and this can cause more harm among platform users.

While conducting research with vulnerable populations, it is important to be knowledgeable and follow existing guidelines (e.g. NSPCC, NIHR Research Governance or GMC) and safeguarding protocols, as well as assessing the impact of this kind of research on the researchers. For example, in our study protocol, we describe which procedure to follow (i.e. contact the clinical governance of the platform for further advice) if the researcher becomes aware of censored posts making reference to suicidal ideation or high risk of self-harm. Ethics as a process allows a dynamic relation between the researcher, the participant and the study protocol, adding flexibility to maintain high ethical standards.

In addition, it is essential to be aware of the potential risk of marginalising certain users through automatic decision-making processes such algorithmic censorship. For example, the threshold applied to establish what constitutes meaningful and appropriate comment may not necessarily reflect a consensual agreement among all stakeholders (i.e. users, professionals, regulators). Lack of consultation and stakeholder involvement can introduce tension between the values of the platform, and the subjective view of users. The co-creation of thresholds and algorithm outcome values becomes crucial to minimise unintentional biases and it also contributes to the processes of trust-building and maintenance towards artificial systems. Accordingly, scholars like Pieter (2011) argue that the relation between the explanations provided to users concerning the algorithm's function is critical in cases of online trust as it provides users with insights regarding how the system works.

Finally, it is important to consider whether some users may not be competent to consent due to additional learning or support needs, emotional distress or severe mental health issues. If unsure, researchers could seek academic experts with experience in both the appropriate methods of social media research and also in the field of psychiatry/psychology and e-health. The academic experts could then act as peer reviewers to ensure that sound principles have been applied and that the findings are based on appropriate research methods. Adopting this course of action and working alongside clinicians and other relevant healthcare practitioners would also serve to maximise dissemination, increase the study visibility and therefore reach to the relevant population. Organisations like Mindtech (http://www.mindtech.org.uk/) bring together healthcare professionals, researchers, industry and the public to develop, implement and evaluate new technologies for mental healthcare. Therefore, multidisciplinary research teams including clinicians, linguists and media experts constitute an essential requirement for projects aiming to appropriately explore online mental health concerns among young people.

The main ethical issues and recommendations that researchers studying youth online mental health communication should consider to ensure their approach is more user-centric are listed in Table 1 below:

Table 1 Ethical issues, suggestions and examples relevant for youth online mental health communication research

Issues	Suggestions	Examples
1. Public-private domain distinction online:	If possible, discuss with the data controllers and online community of the specific DMH service about the level of perceived privacy.	In our case, the group was perceived as closed and registration was required. Live forums were perceived as less private and access to data was granted to researchers. The user-to-professional forum was considered confidential and access to researchers was denied to protect the intimacy of the users and reputation of the platform.
2. Confidentiality and security of online data:	To prevent the likelihood of re-identification personal information (e.g. location, names, etc.) should be omitted. At the time of consent, participants should be informed about dissemination plans (e.g. where and to which potential audience results would be presented).	All data granted by Kooth was already anonymised and stored according to current regulation to maximise security and privacy of the data. Participants were also informed that there is no completely secure interaction online. The suggestion of the University of Nottingham's Research Ethics Committee was followed to illustrate the limits of security on the informed consent document: 'As an online participant in this research, there is always the risk of intrusion by outside agents, i.e., hacking, and therefore the possibility of being identified.'
3. Procedures for obtaining valid consent:	Recruitment strategies and solutions for obtaining consent should be discussed and negotiated with the data controllers and users.	A pop-up window for users over 16 was agreed alongside the data controllers, with a 'Read more' option including clear and accessible information about the study, opportunities to contact the research team, etc. All this information was co-produced a priori with young people not registered at the platform.
4. Procedures for ensuring withdrawal rights and debriefing:	Design your study to allow for the deletion of data points both within the quantitative and qualitative data analysis. Inform the participants about the study results and involve them in the interpretation of results.	Users participating in the study could opt-out at any time by clicking 'I want to withdraw from the study'. A time limit for withdrawal was clearly indicated along with the reason why (e.g. outcomes dissemination). We are planning to send a summary of the findings to the data controllers to be published on the platform.
5. Implications for scientific value and potential harm.	Consider barriers to participation, representativeness of your sample, and whether vulnerable participants are competent to consent. Include a harm-minimisation process for emotional and social harm. Consider implicit algorithmic biases that could discriminate users.	A reference group was created to consult research issues. The research team liaised with the clinical team responsible for moderating the live forum to ensure potential harm to users was identified and communicated in a timely manner.

Table 1 (continued)

Issues	Suggestions	Examples
		Meta-data analysis and stakeholder engagement to detect potential tension between the values of the platform and users as to what constitutes inappropriate content.

3 Conclusions

When planning and designing DMH research to the highest standards of quality, we recommend a process-based approach to ethics (Markham and Buchanan 2012). DMH research is a new area and many of the ethical issues which arise with it are similarly new. Yet this novelty provides opportunities as well as new considerations, and using ethical questions as a way of focussing and improving project design and analysis can help to deliver the maximum potential of each research project. We suggest approaching consent as a process rather than a one-off (i.e. all or nothing) event. This implies an ongoing process that takes into consideration users' autonomy over their personal data, as well as expectations of privacy and intimacy when interacting with the platform to ensure users' rights to privacy are well protected. The development of a stakeholder group—if appropriate—to inform and optimise the process-based approach to ethics becomes mandatory for those researchers interested in championing the Responsible Research and Innovation (RRI) governance framework (Stahl 2013).

RRI helps secure ICT research and innovation are socially accepted and desirable. Our approach to ethics aims to promote children and other 'vulnerable groups' to be more involved in digital research as well as educating the public and gatekeepers on digital citizenship and digital agency. The unique question of censured content has its limitations and, as a next step, this research team is currently working towards the development of a stakeholder group that could inform and support the process of overcoming the ethical issues presented in this paper.

Even though we are still gathering data (i.e. censored posts) within the agreed 12-month window, preliminary results show good levels of participation and plausible suggestions to improve the sensitivity and specificity of the algorithms designed for detecting inappropriate language. Moreover, instead of censoring the whole post, an optimised solution includes blocking only the specific word while reminding the user about the platform Terms of Use (e.g. 'You must be polite on the site in your dealings with counsellors, other workers, and other users'). Censored posts contain sensitive information that should not be automatically deleted as it provides relevant information for assessments and potential referrals to other services.

By ensuring users' rights are respected, we are promoting a culture of e-trust (Taddeo and Floridi 2011), a much needed value in today's research environment which can promote the willingness of the general population to participate in more research projects aiming at societal benefits (Schaefer et al. 2009). Indirectly, a user-centric approach to ethics could also have implications for educating the public—and gatekeepers—by providing them with good exemplars for future practice. For example,

the timely demands for increasing transparency concerning how young people's personal data, and other vulnerable groups, is being used and who is holding and profiting from such data is putting pressure on online platforms to simplify and make more accessible their Terms and Conditions (Coleman et al. 2017). Moreover, the application of the intent of the General Data Protection Regulation when it comes into force in 2018[2] will introduce legislation to improve digital services offered to children and young people. The GDPR will insist that 'where services are offered directly to a child, you must ensure that your privacy notice is written in a clear, plain way that a child will understand' (ICO 2016).

In preparing documentation for any ethical review process, researchers should be explicit about the nature of the online environment and their access to it. These considerations apply not only to mental health data, but to all kinds of data from social media platforms. For example, it is important to state (1) whether the online environment is open, public or private; (2) whether permission has been obtained from the list owner or site administrator to recruit participants or post messages on the site; (3) whether permission to use archive or prospectus data from a list or site has been granted; and finally (4) to describe how subjects will be identified in any reports (e.g. research IDs, usernames, pseudonyms, etc.).

Unfortunately, there are still ethical questions to be resolved regarding this case study. The issue of whether people with mental illness or experiencing mental distress are competent to consent to the study has not been resolved yet. Perhaps the most useful solution to these complex challenges lies with a form of 'negotiated ethics' (Convery and Cox 2012), which would involve seeking advice from the service users themselves as well as from the platform owners/moderators and clinical advisory team who will have valuable experiences and opinions.

Finally, data derived from mental health issues is always intimate and personal, no matter how or where the data was collected. Furthermore, since such data is provided by human participants who are ultimately free to choose whether they want to contribute to research studies or not, the ability to perform such internet-mediated research critically depends on the level of trust that people have in the research community. If people understand how their data is being used and can feel confident about the benefits that the analysis of this data can offer to themselves and society in general, then they will not only willingly contribute their data but may even choose to actively participate in further studies such as citizen science projects. Without transparency of methods, clear ethics guidelines and technical safeguards against (inadvertent) invasions of privacy, public opinion could call for a boycott on internet-mediated research similar to the backlash against genetically manipulated crops that was triggered in the EU in the 1990s (Carr and Levidow 2000). Following in the wake of the explosion in popularity and size of social media services over the last decade, internet-mediated research including web-based questionnaires, social media analysis and web analytics has rapidly risen to become one of the most publically visible forms of social science. At the same time, the prominence of internet related stories in the media means that this kind of research is under heightened scrutiny from the public media.

[2] Although children in the UK are currently covered by the GDPR, this is not guaranteed given the UK's uncertain relationship with the EU following Brexit.

Ethics guidelines and institutional review boards play an important role in establishing an environment of trust, where the public knows what kind of research practices they can expect and researchers can gain confidence in their methods by knowing whom to turn to for an objective evaluation. Due to public concerns over online data security and privacy, thought should be given to the relevant organisational reputation and public trust in research.

Acknowledgements This work was supported by the Economic and Social Research Council [grant number ES/M00161X/1]. Elvira Perez Vallejos acknowledge Jennifer Martin acknowledges the financial support of the NIHR Nottingham Biomedical Research Centre and NIHR MindTech Healthcare Technology Co-operative

References

Adolphs, S. (2006). *Introducing Electronic Text Analysis*. Abingdon: Routledge.

Alderson, P., & Morrow, V. (2004). *Ethics, social research and consulting with children and young people*. Essex: Barnardo's.

American Civil Liberties Union (2015). *Internet Privacy. American Civil Liberties Union*. Available at: www.aclu.org/issues/privacy-technology/internet-privacy.

Anderson, R. (1992). ACM code of ethics and professional conduct. *Communications of the ACM, 35*(5), 94–99.

Article 29 Data Protection Working Party (2011). Opinion 15/2011 on the Definition of Consent. WP187.

Article 29 Data Protection Working Party (2014). Opinion 5/2014 on Anonymisation Techniques. WP216.

Baker, P., Hardie, A., & McEnery, T. (2009). *A Glossary of Corpus Linguistics*. Edinburgh: Edinburgh University Press.

Barbaro, M., & Zeller T. (2006). A face is exposed for AOL searcher no. 4417749, *New York Times*, August 9, 2006.

Bergin, A., & Harding, C. (2016). Ethics and governance in digital mental health research—a joint academic and provider perspective. *Frontiers in Public Health*. **Conference Abstract: 2nd Behaviour Change Conference: Digital Health and Wellbeing**. https://doi.org/10.3389/conf.FPUBH.2016.01.00035.

Berry, D. (2004). Internet research: privacy, ethics and alienation. *Internet Research, 14*, 323–332.

Boyd, D. (2012). The Politics of 'Real Names': power, context, and control in networked publics. *Communications of the ACM, 55*(8), 29–31.

Boyd, D. (2014). *It's Complicated: Social Lives of Networked Teens*. New Haven: Yale University Press.

Boyd, D., & Crawford, K. (2012). Critical questions for big data. *Information, Communication & Society, 15*(5), 662–679.

British Association for Applied Linguistics (BAAL) (2006). Recommendations on Good Practice in Aplied Linguistics. Available at http://www.baal.org.uk/dox/goodpractice_full.pdf.

British National Corpus (BNC XML Edition v3) Distributed by Oxford University Computing Services on behalf of the BNC Consortium. Available: http://www.natcorp.ox.ac.uk/.

British Psychological Society. (2013). *Ethics Guidelines for Internet-mediated Research*. Leicester: British Psychological Society **Available at http://www.bps.org.uk/system/files/Public%20files/inf206-guidelines-for-internet-mediated-research.pdf**.

British Psychological Society (2014). Code of Human Research Ethics. Leicester: UK. Available at http://www.bps.org.uk/sites/default/files/documents/code_of_human_research_ethics.pdf.

Buchanan, E., & Hvizdak, E. (2009). Online survey tools: ethical and methodological concerns of human research ethics committees. *Journal of Empirical Research on Human Research Ethics, 4*(2), 37–48. https://doi.org/10.1525/jer.2009.4.2.37 **Available from: http://internetresearchethics.org**.

Burns, J. M., Davenport, T. A., Durkin, L. A., Luscombe, G. M., & Hickie, I. B. (2010). The internet as a setting for mental health service utilisation by young people. *The Medical Journal of Australia, 192*(11 Suppl), S22–S26.

Carr, S., & Levidow, L. (2000). Exploring the links between science, risk, uncertainty, and ethics in regulatory controversies about genetically modified crops. *Journal of Agricultural and Environmental Ethics, 12*(1), 29–39.

Carter, C. J., Koene, A., Perez Vallejos, E., Statache, R., et al. (2015). Understanding academic attitudes towards the ethical challenges posed by social media research. *ACM SIGCAS Computers and Society, 45*(3), 202–210.

Clarke, L., & Abbott, L. (2015). Young pupils', their teacher's and classroom assisstants' experiences of iPads in a Northen Ireland school: 'four and five years old, who would have thought they could do that?'. *British Journal of Educational Technology*. https://doi.org/10.1111/bjet.12266.

Coleman S., Pothong K., Perez Vallejos E., & Koene A. (2017). The Internet of Trial: How children and young people deliberated about their digital rights. Report available at http://casma.wp.horizon.ac.uk/wp-content/uploads/2016/08/Internet-On-Our-Own-Terms.pdf.

Convery, I., & Cox, D. (2012). A review of research ethics in internet-based research. *Practitioner Research in Higher Education, 6*(1), 50–57 **Retrieved from http://194.81.189.19/ojs/index.php/prhe/article/view/100**.

Corti, L., Van den Eynden, V., Bishop, L., & Wollard, M. (2014). *Managing and Sharing Reserch Data—A Guide to Good Practice*. Thousand Oaks: Sage Publications Ltd..

Data Protection Act (1998). Available at http://www.legislation.gov.uk/ukpga/1998/29/data.pdf.

Denham E. (2016). How the ICO will be supporting the implementation of the GDPR.ICO. Retrieved 3 March 2017 from https://iconewsblog.wordpress.com/2016/10/31/how-theico-will-be-supporting-the-implementation-of-the-gdpr/.

Department of Health (2015). *Report of the work of the Children and Young People's Mental Health Taskforce. Future in mind.* Available at: www.gov.uk/government/publications/improving-mental-health-services-for-young-people.

Dockett, S., & Perry, B. (2011). Researching with young children: Seeking assent. *Child Indicators Research, 4*(2), 231–247.

Economic and Social Research Council (ESRP) (2015). Framework for Research Ethics (FRE). Available at: http://www.esrc.ac.uk/_images/framework-for-research-ethics_tcm8-33470.pdf.

Edwards L. (2016). Brexit: 'You don't know what you've got till it's gone', 13:2 *SCRIPT-Ed* (pp. 112–117). https://script-ed.org/article/brexit-you-dont-know-what-youve-got-till-its-gone/. Accessed 13 September 2017.

Erickson K., Heald P., Homberg F., Kretschmer M., & Mendis D. (2015). *Copyright and the value of the public domain*. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/415014/Copyright_and_the_value_of_the_public_domain.pdf.

Ess C. (2002). Ethical Decision-Making and Internet Research: Recommendations from the AoIR Ethics Working Committee. Available at http://aoir.org/reports/ethics2.pdf.

EU Data Protection Directive 95/46/EC. Available at: http://ec.europa.eu/justice/policies/privacy/docs/95-46-ce/dir1995-46_part1_en.pdf. Accessed 27 September 2017

General Medical Council. Consent to research: research involving vulnerable adults. Available at http://www.gmc-uk.org/guidance/ethical_guidance/6471.asp.

Health Research Authority. Medicines for Human Use (Clinical Trials Regulations) (2004). Informed consent in clinical trials. Available at: http://www.hra.nhs.uk/documents/2014/04/nres-guidance_informed-consent-ctimps_v3-1_2014-04-14.pdf. Accessed 10 September 2017.

Hudson, J. M., & Bruckman, A. (2004). 'Go away': Participant objections to being studied and the ethics of chatroom research. *The Information Society, 20*, 127–139. https://doi.org/10.1080/01972240490423030.

Hunt, D., & Harvey, K. (2015). Health communication and corpus linguistics: using corpus tools to analyse eating disorder discourse online. In P. Baker, & T. McEnery (Eds.), *Corpora and discourse studies: integrating discourse and corpora* (pp. 134–154). London: Palgrave.

ICO. (2012). *Anonymisation: managing data protection risk: code of practice*. Wilmslow: Information Commissioners Office.

ICO. (2014a). *Big data and data protection*. Wilmslow: Information Commissioners Office **https://ico.org.uk/media/for-organisations/documents/1541/big-data-and-data-protection.pdf**.

ICO. (2014b). *Anonymization: managing data protection risk code of practice*. Wilmslow: Information Commissioners Office **https://ico.org.uk/media/for-organisations/documents/1061/anonymisation-code.pdf**.

ICO (2016). Overview of the General Data Protection Regulation (GDPR) accessed via https://ico.org.uk/for-organisations/data-protection-reform/overview-of-the-gdpr/.

Kidd, J., & Finlayson, M. (2006). Navigating unchartered water: research ethics and emotial engagement in human enquiry. *Journal Psychiatric and Mental Health Nursing, 13*, 423–428.

Kramer, Adam. (2014). Facebook posts. Available at: //www.facebook.com/akramer/posts/10152987150867796.

Kramer, A., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences, 111*(24), 8788–8790.

Markham A., & Buchanan E. (2012). Ethical Decision-Making and Internet Research: Recommendations from the AoIR Ethics Working Committee (Version 2.0). Available at http://aoir.org/reports/ethics2.pdf.

McEnery, T., & Wilson, A. (2001). *Corpus Linguistics: An Introduction*. Edinburgh: Edinburgh University Press **(Section 2) ISBN: 0-7486-0808-7**.

Metcalf J. (2016). Human-Subjects Protections and Big Data: Open Questions and Changing Landscapes. *Council for Big Data, Ethics, and Society.* Available at: http://bdes.datasociety.net/council-output/human-subjects-protections-and-big-data-open-questions-and-changing-landscapes/.

NIHR Research Governance. HR good practice resource pack. The research passport: vetting and barring scheme guidance. Available at http://www.nihr.ac.uk/files/Research%20Passport%20Current/Research_Passport_and_the_Vetting_and_Barring_Scheme_Guidance.pdf.

Nissenbaum, H. (2004). Privacy as contextual integrity. *Washington Law Review, 79*(1), 101–158.

Nissenbaum H. (2015). Respecting Context to Protect Privacy: Why Meaning Matters,' *Science and Engineering Ethics*, published online on July 12. Available at http://link.springer.com/article/10.1007%2Fs11948-015-9674-9.

NSPCC. Ethical issues in research with children (resources). Availablea at http://www.nspcc.org.uk/Inform/research/reading_lists/ethical_issues_in_research_with_children_wda55732.html.

Ofcom (2014). The communications market report 7th August 2014. Available at http://stakeholders.ofcom.org.uk/binaries/research/cmr/cmr14/2014_UK_CMR.pdf.

Office for National Statistics (ONS) (2014). Statistical Bulletin. *Internet Access –Households and Individuals* http://www.ons.gov.uk/ons/dcp171778_373584.pdf.

Ohm, P. (2010). Broken promises of privacy: responding to the surprising failure of anonymization. *UCLA Law Review, 57*, 1701.

Parsons, S. (2015). The potential of digital technologies for transforming informed consent practices with children and young people in social research. *Social Inclusion, 3*(6), 56–68.

Perez Vallejos E. (2015). *Lessons to be learned from the Samaritans Radar turmoil.* Blog in CaSMa Research. Available at: www.casma.wp.horizon.ac.uk.

Pieter, W. (2011). Explanation and trust: what to tell the user in security and AI. Ethics and Inf. *Technology, 13*(1), 53–64. https://doi.org/10.1007/s10676-010-9253-3.

Richwood, D. J., Mazzer, K. R., & Telford, N. R. (2015). Social influences on seeking help from mental health services, in-person and online, during adolescence and young adulthood. *BMC Psychiatry, 15*, 40. https://doi.org/10.1186/s12888-015-0429-6.

Robinson, K. M. (2001). Unsolicited narratives from the Internet: a rich source of qualitative data. *Qualitative Health Research, 11*(5), 706–714. https://doi.org/10.1177/104973201129119398.

Schaefer, G. O., Emanuel, E. J., & Wertheimer, A. (2009). The obligation to participate in biomedical research. *JAMA : The Journal of the American Medical Association, 302*(1), 67–72. https://doi.org/10.1001/jama.2009.931.

Sharkey, S., Jones, R., Smithson, J., Hewis, E., Emmens, T., Ford, T., & Owens, C. (2011). Ethical practice in internet research involving vulnerable people: lessons from a self-harm discussion forum study (SharpTalk). *Journal of Medical Ethics, 37*(12), 752–758. https://doi.org/10.1136/medethics-2011-100080.

Stahl, B. C. (2013). Responsible research and innovation: the role of privacy in an emerging framework. *Science and Public Policy., 40*(6), 708–716. https://doi.org/10.1093/scipol/sct067.

Steinfeld, N. (2016). 'I agree to the terms and conditions': (how) do users read privacy policies online? An eye-tracking experiment. *Computers in Human Behavior, 55*, 992–1000 https://doi.org/10.1016/j.chb.2015.09.038.

Stern, S. R. (2003). Encountering distressing information in online research: a consideration of legal and ethical responsibilities. *New Media & Society, 5*(2), 249–266. https://doi.org/10.1177/1461444803005002006.

Taddeo, M., & Floridi, L. (2011). 'The case for E-trust'. Ethics and Inf. *Technology, 13*(1), 1–3. https://doi.org/10.1007/s10676-010-9263-1.

Tait, A., Voepel-Lewis, T., & Malviya, S. (2007). Presenting research information to children: a tale of two methods. *Anesthesia & Analgesia, 105*(2), 358–364.

Tisdall, K., Davis, J., & Gallagher, M. (2009). *Researching with children and young people: research design, methods and analysis*. London: Sage.

Townsend L., & Wallace C. (2016). Social Media Research: A Guide to Ethics. Available at: www.dotrural.ac.uk/socialmediaresearchethics.pdf.

Verma, I. M. (2014). Editorial expression of concern and correction. *Proceedings of the National Academy of Sciences, 111*(29), 10779.

Von Hannover v Germany (2005). 40 European Human Rights Reports 1. Available at https://monash.rl.talis.com/items/CAF8ADD6-8608-3584-A037-D2B2A73C3F03.html.